

# Low-Cost Polarimetric Imaging for Surveillance

Philip Birch, Rupert Young, and Chris Chatwin

Dept. of Engineering and Design, University of Sussex, Falmer, East Sussex, UK

## ABSTRACT

The surveillance industry has traditionally focused on the use of colour intensity images and then used computer vision methods to extract information. Deep learning methods have been demonstrated successfully but require significant computational resources. Fog and rain still present a problem to these methods. Other non-optical imaging technologies are available but the applications can be cost sensitive. Polarimetric cameras offer a solution to some of these problems. This paper presents a practical and low cost design that uses between two and four HD cameras with a wide field of view. This system has an automatic calibration stage that ensures the video frames are synchronised in time. To produce the Stoke parameters each pixel from one camera must be mapped to the others. To perform this, a homography matrix for each camera is automatically discovered and maps each video stream into the correct spatial coordinates. This attempts to use SIFT keypoint mapping but since each input image is a different polarisation state there are potentially a low number of keypoints so an additional check stage is introduced. Calibration results are presented along with example images, post process methods and feature extraction results.

**Keywords:** CCTV, polarimetry, haze reduction, Stokes vector

## 1. INTRODUCTION

Most security cameras only capture the intensity and colour of a scene. Depth clues have been used by implementing ranging cameras, and high dynamic range imaging is useful for images with dark and bright regions. Reference 1 gives an overview of many methods used in detecting pedestrians. There has been considerable progress over the last few years in employing computer vision to segment and understand the images presented. Using modern deep-learning there has been a dramatic increase in the quality of understanding. These systems, however, ignore additional information that can easily be recovered and for a low cost. Imaging of the polarisation states of the light, in addition to the colour and intensity, reveals information about the material, illumination and dielectric properties of objects within the scene.

There are many designs of polarimetric imaging systems. These can be broken down into roughly two types: those that split the amplitude of the beam and make simultaneous measurements of the polarisation state with multiple sensors, and those that temporally multiplex the image with some active polarisation modulator (usually a rotating element or a liquid crystal device). For security applications, we are mainly interested in video data with possibly fast-moving subjects such as pedestrians or motor-vehicles. This makes the temporal modulation methods problematic since the subject may have moved between frames, introducing substantial measurement errors. Therefore, for high-speed applications, the amplitude division methods must be used.

Amplitude division polarimetric cameras usually work by either splitting the light field into a number of different paths. Each path has a different polarisation measurement. The image is either captured by multiple sensors built into one camera or an existing single sensor is sub-divided into different regions. The latter is more compact but the resolution of the camera is reduced by a factor of four (depending on the design).

There are many methods based on the two approaches for measuring the polarisation of imaging. A good review is provided by Ref. 2.

The use of the multiple sensors or active optical systems means that often the design of polarimetric cameras is expensive. It is perhaps this that has reduced its uptake in the security industry apart from specialist applications such as marine surveillance.

There are a number of demonstrated applications of the polarimetric imaging that are of interest. One particular problem that it can help to overcome is that of trying to image in bad weather conditions. Rain and fog can be particularly difficult.

---

Further author information: Send correspondence to Philip Birch E-mail: p.m.birch@sussex.ac.uk

However, the polarised light remains in the same polarisation state after becoming scattered. By processing the image, light from atmospheric interference can be removed. This has been demonstrated by the removal of haze<sup>3</sup> and fog<sup>4,5</sup> and long range imaging through fog.<sup>6</sup> Contrast enhancement through rain has been shown<sup>7</sup> using circularly polarised light.

In addition to this polarisation has been shown to be useful for detecting water. A water detection system was demonstrated<sup>8</sup> for autonomous robots. This could be useful for the detection of pooled water, lakes, etc., in security systems as well as autonomous vehicles.

All reflected light has had its polarisation state modified to some extent. Different dielectric materials will partially polarise light by different amounts. Material identification has been demonstrated by Ref. 9 among others. This has been used to aid the segmentation of objects<sup>10,11</sup> and target detection.<sup>12,13</sup> Disturbed earth has a different scattering property to when compared to non-disturbed and this has been used to detect landmines and improvised explosive devices (IEDs).<sup>14</sup>

Glare reduction and specular illumination estimation can also be estimated. Ref. 15 showed that polarisation can be used to separate Lambertian and specular light. This opens the possibility of using this for colour constancy.

There are therefore a number of benefits to imaging with polarisation but to be viable for the security industry a number of factors have to be considered:

1. The frame rate of the device must be reasonable. Moving objects may cause errors in the measurement of the polarisation state if temporal multiplexing is used.
2. Cost – many existing polarimetric cameras are prohibitively expensive.
3. Light efficiency. The introduction of polarisation elements will reduce the brightness of the image, more so in some designs.
4. Mechanical robustness – cameras with moving parts and delicate optics are problematic, especially if the camera is to be housed outside in a vandal proof box or on a vehicle.

## 1.1 Definition of terms

Since we are dealing with multi-spectral partially polarised light, the polarisation state is better described with a Stokes Vector rather than Jones vectors. We are also not interested and do not know the phase of the light. A polarised light wave, travelling along the  $z$  axis can be described as a sum of two orthogonal electromagnetic waves.

$$\mathbf{E}(z, t) = E_x \hat{\mathbf{x}} + E_y \hat{\mathbf{y}} \quad (1)$$

where

$$E_x = a_x \cos(\omega(t - z/c) + \phi_x) \quad (2)$$

$$E_y = a_y \cos(\omega(t - z/c) + \phi_y) \quad (3)$$

where  $\omega$  is the angular frequency,  $(t - z/c)$  is the time varying component,  $a_x$  and  $a_y$  are the amplitudes, and  $\phi_x$  and  $\phi_y$  are the absolute phases of the wave components. The variables  $a_x$ ,  $a_y$ ,  $\phi_x$  and  $\phi_y$  allow us to define the different linear, elliptical or circular polarisation states. It is convenient to then reformulate these into a Stokes vector  $[s_0, s_1, s_2, s_3]^T$  where

$$s_0 = a_x^2 + a_y^2 \quad (4)$$

$$s_1 = a_x^2 - a_y^2 \quad (5)$$

$$s_2 = 2a_x a_y \cos \phi \quad (6)$$

$$s_3 = 2a_x a_y \sin \phi \quad (7)$$

From this we see  $s_0$  is the total intensity of the beam,  $s_1$  is the difference between the two intensities, and  $s_2$  and  $s_3$  are proportional to the real and imaginary parts of the total wave. From these, a number of useful metrics can be derived. The degree of linear polarisation is the fraction of the light attributed to the linearly polarised part and is:

$$DOLP = \frac{\sqrt{s_1^2 + s_2^2}}{s_0} \quad (8)$$

The degree of circular polarisation is:

$$DOCP = \frac{s_3}{s_0} \quad (9)$$

and the degree of polarisation is the ratio of the total polarised light intensity to the total intensity:

$$DOP = \frac{\sqrt{s_1^2 + s_2^2 + s_3^2}}{s_0} \quad (10)$$

The angle of polarisation is the angle of the elliptical major axis to the  $x$  axis and is given by:

$$\phi = \frac{1}{2} \arctan(s_2/s_1) \quad (11)$$

When the wave is imaged with a conventional camera without any polarisation filters, the pixel level intensity is given by the time average of Equation (1):  $\langle |\mathbf{E}|^2 \rangle = s_0$ . The purpose of the polarimetric camera is to – at least partially – discover the additional variables as these provide information on the two amplitudes and phase differences between the beams and describe these in terms of the Stokes parameters. This can be done by taking a number of intensity measurements,  $I_q$  where  $q$  is either the angle of a linear polariser placed in front of the camera or the left (L) or right (R) hand circular polarisation measurement:

$$\begin{aligned} s_0 &\propto I_0 + I_{90} \\ s_1 &\propto I_0 - I_{90} \\ s_2 &\propto I_{45} - I_{135} \\ s_3 &\propto I_L - I_R \end{aligned} \quad (12)$$

Most sources of light such as sunlight and man-made lighting produce unpolarised light. However, there are several natural processes that tend to polarise light. The major process we need to consider is polarisation by reflection. In this process, different polarisation states are either reflected or transmitted. The amount of this is material dependent. This means by measuring the amount of polarisation reflected from an object we can determine information about the material the object is made from. We can, therefore, segment a road surface from metallic or glass surfaces using the amount of polarisation rather than using the conventional intensity or colour of an object to do this.

De-hazing of images can be achieved by projecting the Stokes vector into a linearly polarised image as below:

$$\mathbf{S}_{\text{corrected}} = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\theta) & 0 & 0 \\ \cos(2\theta) & 1 & 0 & 0 \\ 0 & 0 & \sin(2\theta) & 0 \\ 0 & 0 & 0 & \sin(2\theta) \end{bmatrix} \cdot \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} \quad (13)$$

$\theta$  is discovered by optimising for the contrast ratio. An example is shown in Figure 1 which is a telephoto image of a wind farm construction site in the English Channel. The de-hazed image shows a considerable improvement. This was captured using a Canon D50 SLR camera with a manually rotated linear polariser attached. Other examples images show the usefulness of the method for surveillance.

Figure 2 shows the Stokes parameter images for the three channels that were measured for an image of a car-park and lighthouse. The images are converted to grey scale before conversion although each colour channel could be separately processed.

In Figure 3 the images have been further processed. The phase angles and the DOLP are shown. The natural log of the DOLP is also shown to emphasize the differences. Here, the difference between the metallic objects, brick objects and grass can clearly be seen. Figure 3 d) shows a zoomed in section of the Laplacian of the Gaussian of the DOLP. From this, the metallic and glass objects (cars, street furniture, parts of some buildings) can clearly be located.

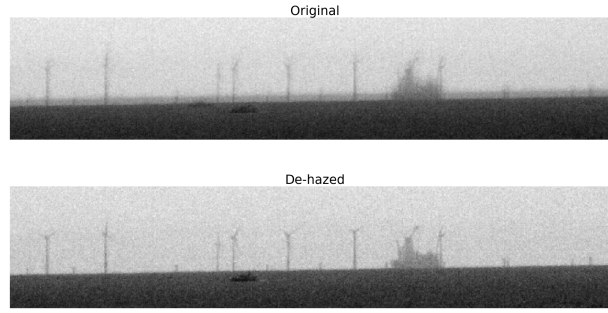


Figure 1: Example of de-hazed image. The top shows the original image and the lower one is the de-hazed. The images are of the construction of a offshore wind farm.

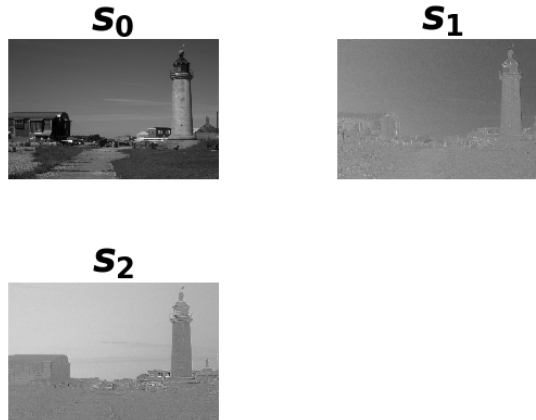


Figure 2: The three Stokes parameter images.

## 2. POLARISATION MEASUREMENT

The measurement of the Stokes parameters can be achieved by taking a number of measurements with different sets of polarisation filters. It is, however, complicated by the wavelength dependency of the polarisation and the difficulty of measuring the circular states. For this reason, many polarimeters only measure some of the Stokes parameters and it is still possible to determine useful information from this. The alternatives are:

1. One-dimensional polarimetry – A single measurement is made with a fixed linear polariser. Commonly used in photography, it can be used to remove reflection from specular surfaces such as water.
2. Two-dimensional polarimetry – Two different polarisation measurements are determined. With this  $s_0$  and  $s_1$  can be determined. This has been useful for clutter rejection and separating objects of interest from the background.
3. Three-dimensional polarimetry – Multiple linear polarisation states are measured to obtain  $s_0$ ,  $s_1$  and  $s_2$ . From this the DOLP and  $\phi$  can be determined.
4. Full Stokes polarimetry – this measures all four parameters. The difficulty is measuring the circular state, which requires retardance plates making it wavelength dependent. For this reason, many authors only implement the three-dimensional polarimeter.

## 3. CAMERA DESIGN

In this paper, we will explore a two and three dimensional polarimetric design. This is based on keeping the cost low, whilst still extracting useful information. There are many commercial designs that use a single camera but the requirement

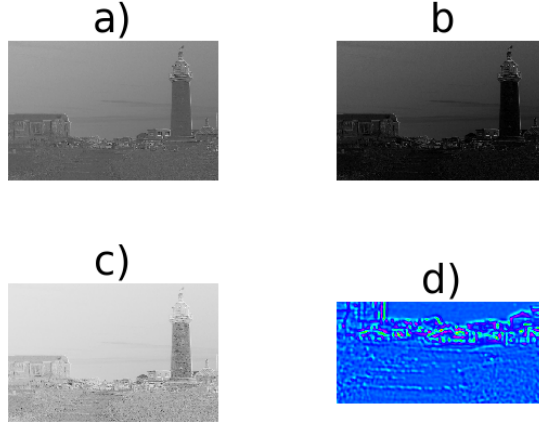


Figure 3: Processed images: a) Angle ( $\phi$ ); b) DOLP image; c) the natural log of the DOLP; d) the Laplacian of Gaussian of the DOLP (false colour)

to run video means that there must be a variable filter and the capability to take multiple shots per frame. We therefore implemented a custom design using either two or four separate low-cost cameras. The cameras are Mobius ActionCams that can record at 1080p ( $1920 \times 1800$ ) at 60 fps. The video is compressed using H264 AVC with a variable bit rate. The images, with audio, can be sent directly to a host computer via USB or alternatively stored on an internal SD card. The cameras were set up to record to SD card as soon as they were powered on. The layout of the cameras is shown in Figure 4.

Each camera has a fixed focus with a fixed linear polariser. It can be shown<sup>16,17</sup> that the optimum angles for the polarisers for  $N$  cameras is

$$\theta_i = \frac{i-1}{N}\pi \quad (14)$$

and for  $N = 4$  these values are at  $0^\circ, 45^\circ, 90^\circ$  and  $135^\circ$  in front for the four camera case and  $0^\circ$  and  $90^\circ$  for the two camera,  $N = 2$ , case. Each camera was individually calibrated using the method described in Ref. 18 to correct for the intrinsic distortion. The cameras had a significant amount of barrel distortion due to the wide angle of the lens. The image was then warped to a flat image. Due to the amount of distortion, this sometimes had the effect of reducing the effective field of view so this step was optional.

Since the videos are not synchronised there is a certain random frame difference that occurs between each camera. The cross-correlation of the audio tracks was then taken to find the frame difference. A window of  $W$  audio samples is extracted from the left-hand channel at the beginning of the  $j^{th}$  video file  $A_j$ :

$$O_j = \underset{t}{\operatorname{argmax}} \mathcal{F}^{-1} \{ \mathcal{F}(A_1[t]) \mathcal{F}(A_i[t])^* \} \quad (15)$$

where  $i = 2, \dots, N$  and  $\mathcal{F}$  is the Fourier transform operator. The video synchronisation is then performed by skipping  $O_j - \min(O_i)$  frames for each corresponding stream.

The accuracy of the above calculation is only  $\pm \frac{1}{2}$  time steps. However, the audio is recorded at  $44.1kHz$  and the video a  $25Hz$ . This means using this method, the video can be synchronised to  $\pm 13\mu s$  or  $0.057\%$  of a video frame. This means that two videos should stay within synchronisation for about  $70s$ . To improve on this, there are two options. The FFTs used to calculate Equation 15 could be zero padded to give an arbitrary precision - although in reality this will become limited by machine precision or signal noise. To improve the calculation time and reduced the RAM overheads other methods could be adapted to perform this<sup>19</sup> if needed. The other alternative is to recalculate Equation 15 every few seconds to ensure that the videos remain in synchronisation. The later was used in practice.

The code to perform this is written in Python. OpenCV is used for most of the image processing tasks in this project, but it not suitable for audio. To extract the audio, FFmpeg is used to convert the videos to a set of PCM audio arrays.

FFmpeg also provides information on the audio sample rate and video frame rate of each file. OpenCV is then used to read the image data from a video file.

For the live streaming over USB, the video synchronisation is not needed. Each camera can be captured using OpenCV but in practice the library does not name the video source and the correct camera cannot be guaranteed to be captured in the correct order so FFmpeg is used to pipe the images into Python.

For the video, a prime camera is selected which is the one nearest to the systems optical axis. From this homography projection matrices are discovered.

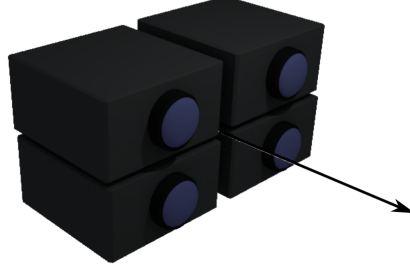


Figure 4: The arrangement of the four cameras. The overall optical axis is shown.

### 3.1 Homography Discovery

A number of automated methods have been presented in the past in which key-points are matched in each image to discover the homography between two cameras. The key-points are often discovered using SIFT or SURF. In this case, SIFT was used. To match the features, each is compared against every over using  $l^2$ -norm, and the closest two returned. Lowe's ratio test is then performed to rule out more false matches. Least squares fitting and RANSAC are used to select a good candidate homography matrix,  $\mathbf{H}_{i,j}$ , that maps from camera  $i$  to camera  $j$ .

However, these methods were not originally designed to work on polarised images. If a highly polarising material is present in the image, there may well be no matching key-points, or there may be regions of the image sets with little information leading to a poor matrix inversion and incorrect  $\mathbf{H}_{i,j}$ . When run on an arbitrary image the matching processed failed a significant number of times and so multiple frames were needed. For this reason, an additional step was introduced. Since the corrected images should have no projective, scale or rotation difference between them, cross-correlation provides a good solution. Due to the size of the images the Fourier transform,  $\mathcal{F}$ , method was used:

$$C = \max |\mathcal{F}^{-1} \{D \cdot \mathcal{F}(I_i) \cdot \mathcal{F}(I_j)^*\}|^2 \quad (16)$$

where  $I$  are the mean subtracted test images and  $D$  is the difference of Gaussian operator in the Fourier domain. If  $C$  is above a threshold, the corresponding  $\mathbf{H}_{i,j}$  is kept. After a number have been collected, the mean of each element of  $\mathbf{H}_{i,j}$  is used to generate a good potential homography matrix. The variance of the elements should be low if there are good matrices, if not the threshold on  $C$  is increased and the process re-run. We note that the variance methods is only indicative of a good potential homography matrix. The values of  $H$  are not independent. In practice, though a low variance did indicate a reasonable estimate and was a simple way to show if the SIFT matching process completely failed.

Once a good set of homography matrices and the frame synchronisation have been determined, the Stokes vector can be calculated via Equation (12). Since we are not measuring circular polarisation  $s_3$  is ignored.

## 4. RESULTS

The camera design at present has a number of limitations. Since each camera does not share an optical axis there is a parallax problem, objects near to the camera will appear in different places in each image. This will result in incorrect Stokes vectors. To overcome this we reduce the resolution of the image to  $640 \times 480$  pixels and maintain a minimum working distance of about  $5m$ . The other restriction is also related to this. In this case, there is an error introduced because of the imperfections in the calculation of the homography matrices and camera calibration. The effect of this is to reduce the useful field of view of the camera. Work is continuing to improve upon this.

#### 4.1 Two channel polarimetric images

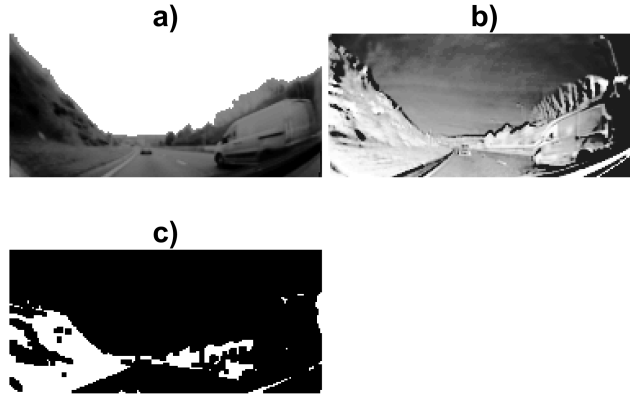


Figure 5: a) Intensity image taken with the two channel camera of a road and verge; b) the DOLP; c) the processed image showing vegetation segmented.

Figure 5 shows an example output from the polarimetric camera running in 2 channel mode showing the image of a road. The DOLP is calculated and processed using simple value thresholds and variation metric:

$$\Lambda = \sum_{i,j} \Gamma(i,j) |D_{i,j} - \bar{D}_{i,j}| \quad (17)$$

where  $\Gamma(i,j)$  is a window function,  $D_{i,j}$  is the DOLP at  $i,j$  and  $\bar{D}_{i,j}$  is the local mean of  $D$ . The vehicles in the scene have a high degree of polarisation but a low local variance meaning they can be segmented from the vegetation easily.

#### 4.2 Three channel polarimetric images

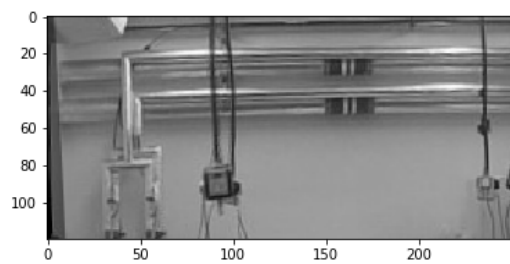
Figure 6 shows a three dimensional result taken with four cameras combined. The metallic objects show a higher degree of polarisation as indicated in the DOLP image.

### 5. CONCLUSIONS

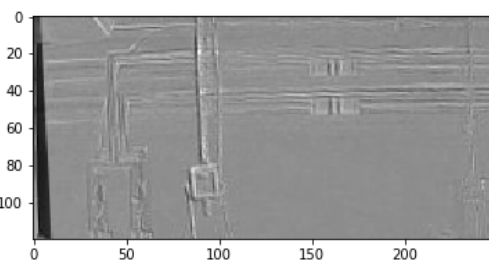
Polarimetric imagery has been neglected in the past due to the cost, difficulty, and robustness of the cameras. It does however, offer more information than an intensity image alone can detect. It can be used to separate materials that may look identical in the intensity image. This can be used to implement an enhanced version of vehicle and pedestrian trackers. This paper has demonstrated how to generate a low-cost robust polarimetric camera using either two or four identical low-cost cameras. Simple processing methods have been used to show how the system can be used to remove haze, segment roads, traffic signs and vegetation. Further work is currently being undertaken to gather larger datasets to experiment with more advanced computer vision methods, such as, deep learning to further explore the advantages of polarimetric imaging.

### REFERENCES

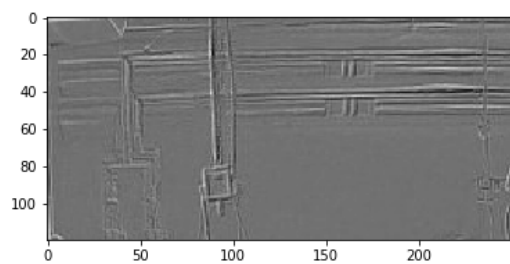
- [1] D. Geronimo, A. M. Lopez, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1239–1258, July 2010.
- [2] J. Tyo, D. Goldstein, D. Chenault, and J. Shaw, "Review of passive imaging polarimetry for remote sensing applications," *Applied Optics*, vol. 45, no. 22, p. 5453, 2006.
- [3] S. K. Nayar, S. G. Narasimhan, and Y. Y. Schechner, "Polarization-based vision through haze," *Applied Optics*, vol. 42, pp. 511–525, Jan. 2003.
- [4] N. Hautiere, J.-P. Tarel, and D. Aubert, "Towards Fog-Free In-Vehicle Vision Systems through Contrast Restoration," in *Proceedings of the IEEE*, (Minneapolis, MN, USA), IEEE, June 2007.



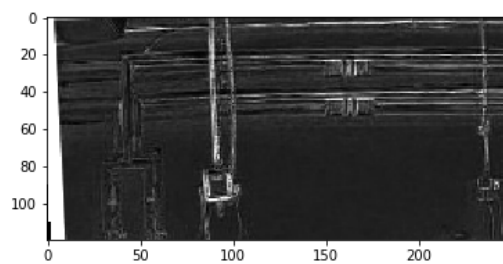
(a)  $S_0$



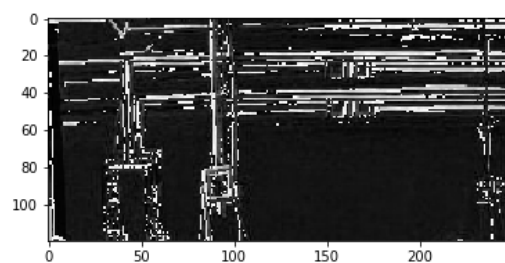
(b)  $S_1$



(c)  $S_2$



(d)  $DOLP$



(e)  $\phi$

Figure 6: 4 camera Stokes image.



- [5] E. Qu, H. Ju, J. Liang, L. Ren, W. Zhang, Y. Tang, Z. Bai, and Z. Wu, "Real-time image haze removal using an aperture-division polarimetric camera," *Applied Optics*, vol. 56, pp. 942–947, Feb. 2017.
- [6] J. Fade, S. Panigrahi, A. Carré, L. Frein, C. Hamel, F. Bretenaker, H. Ramachandran, and M. Alouini, "Long-range polarimetric imaging through fog.," *Applied Optics*, vol. 53, pp. 3854–3865, June 2014.
- [7] F. Liu, J. Xu, P. Han, and X. Shao, "Design of a circular polarization imager for contrast enhancement in rainy conditions," *Applied Optics*, vol. 55, pp. 9242–9249, Nov. 2016.
- [8] B. Xie, Z. Xiang, H. Pan, and J. Liu, "Polarization-based water hazards detection for autonomous off-road navigation," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 3186–3190, 2007.
- [9] S. Tominaga and A. Kimachi, "Polarization imaging for material classification," *Optical Engineering*, vol. 47, no. 12, p. 123201, 2008.
- [10] Y. Zhao, L. Zhang, D. Zhang, and Q. Pan, "Object separation by polarimetric and spectral imagery fusion," *Computer Vision And Image Understanding*, vol. 113, no. 8, pp. 855–866, 2009.
- [11] Y.-Q. Zhao, P. Gong, and Q. Pan, "Object Detection by Spectropolarimetric Imagery Fusion," *Ieee Transactions On Geoscience And Remote Sensing*, vol. 46, no. 10, pp. 3337–3345, 2008.
- [12] F. Goudail and P. Réfrégier, "Statistical techniques for target detection in polarization diversity images," *Optics Letters*, vol. 26, no. 9, pp. 644–646, 2001.
- [13] L. Meng and J. P. Kerekes, "Adaptive target detection with a polarization-sensitive optical system," *Applied Optics*, vol. 50, no. 13, pp. 1925–1932, 2011.
- [14] F. Cremer, W. de Jong, K. Schutte, W.-J. Liao, and B. A. Baertlein, "Detectability of surface-laid landmines with a polarimetric IR sensor," in *Proc. SPIE , Detection and Remediation Technologies for Mines and Minelike Targets VIII*, (Orlando), International Society for Optics and Photonics, Sept. 2003.
- [15] S. Nayar, X. Fang, and T. Boult, "Separation of reflection Components using color and polarization," *International Journal Of Computer Vision*, vol. 21, no. 3, pp. 163–186, 1997.
- [16] J. S. Tyo, "Optimum linear combination strategy for an N-channel polarization-sensitive imaging or vision system," *JOSA A*, vol. 15, pp. 359–366, Feb. 1998.
- [17] F. Goudail and A. Beniere, "Estimation precision of the degree of linear polarization and of the angle of polarization in the presence of different sources of noise," *Applied Optics*, vol. 49, no. 4, pp. 683–693, 2010.
- [18] Zhang, Zhengyou, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [19] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.*, vol. 33, pp. 156–158, Jan 2008.